A peer-reviewed open-access journal
NeoBiota
Advancing research on alien species and biological invasions

NeoBiota 107: 131–153 (2026)
DOI: 10.3897/neobiota.107.176004

Research Article

Temporal genomics unravels demographic and selective invasion dynamics of *Drosophila suzukii* at a European range edge

João Carvalho[1,2], João Pedro Marques[1,2], Sara Sario[3,4], Liliana Farelo[1,2], Conceição Santos[3,4], José Melo-Ferreira[1,2,3]

1 *CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Universidade do Porto, Vairão, Portugal*
2 *BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, Vairão, Portugal*
3 *Departamento de Biologia, Faculdade de Ciências, Universidade do Porto, Porto, Portugal*
4 *LAQV-REQUIMTE, Faculty of Sciences, University of Porto, Porto, Portugal*
Corresponding author: José Melo-Ferreira (jmeloferreira@cibio.up.pt)

Abstract

Drosophila suzukii is a globally invasive fruit fly that causes significant economic impacts on soft-skinned fruit crops. While its invasion history has been studied on a global scale, regional-scale demographic and selective dynamics remain less understood. Portugal is at the westernmost limit of the European expansion, a key region to investigate genetic diversity and adaptation at the most recently colonized areas. Previous whole-genome sequencing of Portuguese populations suggested a Mediterranean invasion route and identified candidate variants potentially involved in local adaptation. Here, we re-sample the Portuguese populations two to four years later, to examine temporal changes in the invasion dynamics. Our results confirm the Mediterranean invasion route and show that the extremely low genetic diversity in one early-sampled population (PT-VM19) reflects a founder event. We show that this population gained genetic diversity over time, likely through gene flow from neighboring populations. We detect signatures of selection along the genome and find that most overlap with previously reported candidate genes is driven by PT-VM19, suggesting that many earlier signals were false positives associated with its low diversity. Yet, nine genes show consistent signatures of selection over time in Portuguese populations, identifying them as potential candidates for local adaptation. These genes are involved in neural development, embryogenesis, metabolism, and chromatin organization. Based on these, we developed and validated a set of SNP markers for monitoring *D. suzukii*. Our study underscores how temporal genomic data from recently invaded regions can uncover demographic recovery and ongoing adaptation and contributes genomic tools for applied population monitoring.

Key words: Invasive pests, local adaptation, pool-sequencing, SNP validation, Spotted wing drosophila, time-series genomics

OPEN ACCESS

Academic editor: Robert Colautti
Received: 24 October 2025
Accepted: 19 February 2026
Published: 4 May 2026

Citation: Carvalho J, Marques JP, Sario S, Farelo L, Santos C, Melo-Ferreira J (2026) Temporal genomics unravels demographic and selective invasion dynamics of *Drosophila suzukii* at a European range edge. NeoBiota 107: 131–153. https://doi.org/10.3897/neobiota.107.176004

Copyright: © João Carvalho et al.
This is an open access article distributed under terms of the Creative Commons Attribution License (Attribution 4.0 International – CC BY 4.0).

Introduction

The invasive spotted wing drosophila, *Drosophila suzukii* Matsumura (Diptera, Drosophilidae), is a highly polyphagous fly, believed to be native to Southeast Asia, though this has been debated (Walsh et al. 2011). Its adaptability, flexible metabolism, and high dispersal capacity, facilitated by the transport of immature stages and increased international trade (Shearer et al. 2016; Walton et al. 2016; De Ros

et al. 2020; Knapp et al. 2020), have enabled its expansion far beyond its presumed native range. The species was first detected outside this range in Hawaii in 1980 (Hauser 2011) and subsequently invaded Europe and North America almost simultaneously in 2008 (Calabria et al. 2010; Rota-Stabelli et al. 2013). Since then, it has rapidly spread across both continents and is now recognized as a major pest, causing substantial damage to economically important soft and thin-skinned fruit crops (Boughdad et al. 2020; Reyes and Lira-Noriega 2020; Garcia et al. 2022).

Genetic studies suggest that despite occurring roughly at the same time, the North American and European invasions are independent demographic events. Analysis of X-linked gene fragments found evidence of differentiation between the European, Asian, and North American populations but detected no differentiation within the U.S. populations (Adrion et al. 2014). Additionally, Fraimout et al. (2017), using microsatellites, identified multiple invasions from Asia into Europe and the U.S., as well as differentiation between the eastern and western U.S. populations. Together, these studies highlight distinct global invasion patterns. North American populations (also including South American, Brazil) and European populations (including La Réunion Island) represent separate invasion routes, each associated with different native source populations and multiple introduction events (Fraimout et al. 2017). In Europe, the species was first detected in Italy and Spain in 2008–2009 (Calabria et al. 2010), from where it rapidly spread along the Mediterranean basin and into central and northern Europe within a few years (Cini et al. 2014). In Portugal, *D. suzukii* was first recorded in 2012, when adults were found in a raspberry greenhouse on the Alentejo coast, in southern Portugal (OEPP/EPPO 2012). Subsequent reports documented its rapid northward spread to other mainland regions and its eventual arrival in the Azores and Madeira archipelagos (Rego et al. 2017; Macedo et al. 2023). A recent whole genome pool sequence survey of three Portuguese populations of the species suggested that the most likely origin of these populations was a southern European invasion route (Sario et al. 2024). Their analyses of genome-wide allele frequency patterns and inferred population relationships placed Portuguese samples closer to those from Spain, Italy, and Southern France than to more northern European populations, supporting an introduction through the Mediterranean coast (Sario et al. 2024). However, the genetic patterns observed in one of the populations did not fully align with this scenario, suggesting potential heterogeneity and complexity in the invasion dynamics of *D. suzukii* in Portugal. This result underscored the need for further analysis to investigate the causes of the low genetic diversity in that population and to clarify its relationship with the other populations.

Given its significant economic impact (Bolda et al. 2010; Yeh et al. 2020), it is important to manage and control invasive *D. suzukii* populations. Effective monitoring of invasive species requires a thorough understanding of the ecological and evolutionary processes that drive their spread and establishment. In this context, tracking invasive populations across different time points allows us to quantify their success and analyze their adaptive potential. Additionally, biological invasions are of broader scientific interest as they serve as natural experiments, offering a unique opportunity to study evolutionary processes on contemporary timescales (Lee and Gelembiuk 2008; Colautti and Lau 2016; Dlugosch et al. 2016). Invasive populations often face differences in biotic interactions, climate, resource availability, and disturbance regimes compared to their native regions. Thus, adaptive evolutionary shifts in response to new selection pressures are likely

crucial for the initial establishment and spread of invasive species following their introduction (Facon et al. 2011; Colautti and Barrett 2013; Ochocki and Miller 2017). Monitoring invasive populations over time offers a promising approach to studying these evolutionary shifts, allowing researchers to track the frequency and dynamics of adaptive variation in natural populations. Despite the growth in theoretical and empirical research on the evolutionary biology of invasive species over the past decade, the genetic basis of evolutionary adaptations during invasions remains largely unexplored (Prentis et al. 2008; Barrett 2015; Reznick et al. 2018), with even fewer studies tracking the same populations across different time points.

Insights into putative adaptive alleles have vast potential applications across various fields, including biomedicine (Hill et al. 2020; Natterson-Horowitz et al. 2023), animal and plant breeding (Daetwyler et al. 2014; Gaur et al. 2018; Schneider et al. 2024) and pest management. Identifying adaptive variation often relies on identifying patterns of sequence variation that are not compatible with the null hypothesis of genetic drift. One of the clearest signatures of selection is generated by the sweep of a beneficial allele through a population. In this context, linked neutral alleles hitchhike along with the selected allele (Smith and Haigh 1974; Barton 2000), leading to characteristic patterns such as reduced genetic variability, increased linkage disequilibrium, and skewed site frequency spectra around the targeted genomic region (Andolfatto 2001; Stephan et al. 2006; Min et al. 2022). These patterns provide inference power for widely used tests for selection (e.g., Nielsen et al. 2005; Foll and Gaggiotti 2008; Pavlidis et al. 2013). Typically, these tests rely on polymorphism data from a single time point. However, analyzing samples collected across multiple time points can increase detection power by leveraging allele frequency changes over time, enabling the inference of selection at a single site. The advantages of time-series data for population genetic analyses have been acknowledged for decades (Fisher and Ford 1947; Wright 1948), and this approach has been extensively applied in experimental evolution studies, particularly those using pooled sequencing of individuals (Pool-Seq; Futschik and Schlötterer 2010; Schlötterer et al. 2014; Taus et al. 2017). Despite these advancements, time-series data remain underutilized in studies of natural populations, although some successful examples exist in European human populations (Mathieson et al. 2015), and in *Drosophila melanogaster* (Bergland et al. 2014).

Here, we study Portuguese populations of *D. suzukii*, which represent the westernmost edge of the European range and a recently invaded area. Analyzing genome-wide polymorphism data collected at two time points separated by two to four years, our specific goals were to: (i) assess rapid temporal demographic changes, (ii) re-assess signatures of selection and the stability and recurrence of these signals over time, and (iii) identify SNP markers suitable for regional monitoring of the invasive populations.

Methods

Sample collection and genomic DNA extraction

Sample collection and DNA extraction were carried out following the protocol described by Sario et al. (2024). In summary, 240 adult *Drosophila suzukii* were collected in 2023 from three locations in northern Portugal: Vieira do Minho (PT-VM23), Oliveira de Azeméis (PT-OAZ23), and Castelo de Paiva (PT-CP23).

Permission to collect fruit and access orchards was granted by landowners, who also assisted in the sampling process. To ensure accurate species identification and prevent contamination with other *Drosophila* species, adult flies were identified based on morphological traits using the European Plant Protection Organization (EPPO) dichotomous key (OEPP/EPPO 2013). Identified specimens were then surface-sterilized with a 1.5% bleach solution (v/v in Milli-Q H_2O) and stored at -80 °C until DNA extraction. Further details on sampling locations, collection years, and methodologies are available in Suppl. material 2: table S1.

Genomic DNA was extracted from sets of thoraxes of 16 adult individuals coming from the same sampling locality (8 females and 8 males). DNA extraction followed the QIAamp DNA Micro Kit (Qiagen) protocol, with minor adjustments following Sario et al. (2024): samples were mechanically disrupted in Buffer ATL using a homogenizer with a ceramic bead prior to proteinase K digestion; the lysis step was extended to 4 hours at 56 °C; and carrier RNA was added after incubation according to the manufacturer's instructions. For each sample/locality, this process was repeated for 5 sets of 16 different individuals, totaling 80 individuals per population. This approach was used to i) increase DNA yield given the small body size of *D. suzukii*, and ii) to ensure balanced representation of both sexes. The five resulting extracts were subsequently combined to generate a single pooled DNA sample per population to prepare the genomic DNA libraries.

This pooling strategy, combined with coverage-aware filtering in downstream analyses (see below), aimed at reducing biases arising from unequal individual contribution or sequencing depth variation among samples. While minor differences in coverage among individuals may remain, these were accounted for in the calculation of allele frequencies and population-genomic metrics.

Library preparation, pool-sequencing, and data processing

DNA libraries were prepared from pooled DNAs (one library preparation per each of the three sampled populations) and sequenced following the protocol described in Sario et al. (2024), using the TruSeq DNA PCR-Free kit (Illumina), with library quantification by quantitative PCR, quality assessment on a 2200 TapeStation (Agilent), and paired-end 150 bp sequencing performed on an Illumina NovaSeq at Novogene Co. Ltd. (Cambridge, UK). Our dataset was expanded by incorporating previously sequenced Portuguese samples from Sario et al. (2024), sampled in 2019 and 2021, alongside a worldwide dataset retrieved from Olazcuaga et al. (2020) (Fig. 1a, b; https://www.ncbi.nlm.nih.gov/bioproject/PRJNA576997). All raw sequencing data were processed according to the protocol outlined in Sario et al. (2024), with parameter details specified therein. In brief, adapter sequences were removed, and reads were quality-trimmed using Trimmomatic v0.39 (Bolger et al. 2014). The resulting paired-end reads were aligned to the *D. suzukii* Dsuz_RU_1.0 reference genome (NCBI RefSeq assembly GCF_037355615.1; analyzed in Ratiu et al. 2024) using BWA-MEM 0.7.18 with default settings (Li and Durbin 2009). The genome assembly is highly complete and contiguous, with a scaffold N50 of 26.7 Mb, contig N50 of 26.7 Mb, and a BUSCO completeness score of 98.8%. Duplicate reads were eliminated using Picard v3.3.0 (http://broadinstitute.github.io/picard/), and those with a mapping quality below 20 (-q 20) were filtered out with SAMtools v1.21 (Li et al. 2009). Mapping statistics were evaluated using bamtools (Barnett et al. 2011), and per-site coverage was obtained

with SAMtools v1.21. All populations showed consistently high mapping quality, with ≥ 99% of reads mapped, > 95% proper pairs, and very low proportions of singleton reads (<1%), indicating that data quality was comparable across populations. A combined mpileup file including all populations and the reference genome was generated with SAMtools v1.21, and variants were called using VarScan v2.4.6 (Koboldt et al. 2012). Finally, reads mapped to autosomal (A) and X chromosome (X) scaffolds were extracted from the VCF file using vcftools (Danecek et al. 2011).

Assessment of population genetic diversity

Genome-wide nucleotide diversity (π) and Watterson's θ were estimated for all 6 focal populations using PoPoolation v1.2.3 (Kofler et al. 2011). Indexed BAM files were used to generate separate datasets for autosomes (A) and the X chromosome (X) with SAMtools v1.21. Mpileup files were then created using SAMtools v1.21, applying quality filters (-q 20, -Q 20). Diversity indices (π and θ) were calculated in non-overlapping 20 kb windows using population- and chromosome-specific filters to account for variation in coverage. These include a minimum allele count of 3, minimum and maximum coverage thresholds of one-half and three times the median coverage, and a minimum covered-site fraction of 0.50, following Sario et al. (2024). Detailed population-specific parameters are provided in Suppl. material 2: table S1. To further illustrate the effect of these filters and the coverage distribution in our dataset, frequency histograms for each population are provided in the Suppl. material 1: figs S1, S2.

Analysis of population structure

VCF files for autosomal (A) and X-linked (X) scaffolds were processed in R using the poolfstat package v2.1.1 (Gautier et al. 2021) to generate a pooldata object. Only sites with a minimum of four reads per pool, coverage below three times the median, and a minor allele frequency (MAF) of ≥ 0.01 were retained. The minimum read count threshold refers to the requirement that each pool must have at least four reads covering a given site for that position to be considered; sites with lower coverage in any pool were excluded. This filter is intended to minimize the inclusion of potential sequencing artefacts rather than to define the coverage used for allele frequency estimation. No minimum read count per allele (parameter *min.rc* in poolfstat) was applied, so alleles supported by few reads were still considered, although indels were excluded. The pooldata object was then converted into a BayPass input file using poolfstat, with a subsample size of 75,000 selected via the thinning method (Gautier et al. 2021). BayPass v2.3 (Gautier 2015) was employed to infer population structure and relationships by estimating a scaled covariance matrix (Ω) of allele frequencies. Population structure was visualized using a heatmap with clustering analysis and singular value decomposition (SVD) of the Ω matrix (Gautier 2015). Additionally, pairwise F_{ST} values between populations were calculated from the pooldata object using poolfstat v2.1.1 with default settings (Gautier et al. 2021).

Genetic relationships and migration patterns were further examined using TreeMix v1.13 (Pickrell and Pritchard 2012), based on allele frequencies subsampled with poolfstat. To evaluate confidence in the inferred tree topology, bootstrap replicates (-bootstrap) were generated without assigning a root. An initial

maximum-likelihood tree was constructed, followed by the sequential addition of migration events (-m), iterated 100 times for each value of *m* (ranging from 1 to 10). The optimal number of migration events was determined using OptM (Fitak 2021).

Detection of selective sweeps

Putative selective sweeps in the focal populations were identified separately for autosomal and X chromosome scaffolds using Pool-hmm v1.4.4 (Boitard et al. 2013), which employs a hidden Markov model approach. Analyses were performed independently for each population, and selection signals were subsequently compared across sampling years within the same location to explore temporal dynamics. All analyses followed the parameter from Sario et al. (2024), with the number of chromosomes in the pool (-n) and minimum (-c) and maximum (-C) coverage thresholds matching those used for estimating diversity indices. Genome-wide *theta* (-t) was derived from PoPoolation estimates, using the mean θ for each population, separately for autosomes and the X chromosome. Our primary focus was on genomic regions classified as under selection, particularly those fragments with a selection probability of 1.

For each genomic region with a selection probability of 1, we considered the region itself along with 5 kb of flanking sequence on either side, and retrieved gene annotations from the reference genome using bedtools intersect. RefSeq IDs were converted into FlyBase gene IDs using FlyBase converting tools (vFB2023_06, released December 12, 2023) (Jenkins et al. 2022). Functional enrichment analysis was conducted with the g:Profiler web server (Raudvere et al. 2019) to identify genes under selection. Only genes within retained segment windows were included as the background gene set. Categories with fewer than five genes were excluded, and multiple testing correction was applied using the g:SCS algorithm (Raudvere et al. 2019). Gene Ontology (GO) term annotation was based on *Drosophila melanogaster*.

Primer design and validation

To facilitate future monitoring of the Portuguese *D. suzukii* invasion and assess its dynamics and success, we designed primers targeting candidate regions inferred to be under selection. We focused on genomic regions with a selection probability of 1 that were shared across all six focal populations, spanning the three locations and two time points. The overlap of predicted selective sweep windows revealed 40 common genes, of which nine had known functions. From these, we selected six genes with relevant biological functions (*Muc68Ca*, *Sema1a*, *Acer*, *H2.0*, *Gpdh1*, and *Uba4*) for downstream SNP validation. To identify target SNPs, we used major and minor allele read counts from PoPoolation2 output and overlaid them with Pool-hmm v1.4.4 selection probabilities. SNPs were retained if (1) they were polymorphic and of sufficient quality (as reported by Popoolation2), and (2) consistently classified as under selection in all focal populations. The sizes of the six selected genes, along with the number of polymorphic sites and target SNPs identified in each, are provided in Suppl. material 2: table S2.

To validate the selected SNPs and assess the suitability of the associated loci for future genetic monitoring, we designed and tested primer pairs for nine amplicons across six target genes (Suppl. material 2: table S3). PCR amplifications were per-

formed using genomic DNA extracted from two individuals (a male and a female), using the QIAamp DNA Micro Kit (Qiagen). PCRs were conducted with QIAGEN Multiplex PCR Master Mix and optimized individually for each primer pair. Thermal cycling profiles generally included an initial denaturation at 95 °C for 15 minutes, followed by 35 cycles of 95 °C for 30 seconds, annealing at 55–69 °C for 20 seconds, and elongation at 72 °C for 20 seconds, with a final extension at 72 °C for 5 minutes. Amplification success was checked via 2% agarose gel electrophoresis. PCR products were purified using Exonuclease I (New England Biolabs) and FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific), then sequenced with an ABI 3500xl Genetic Analyzer.

Results

Genome-wide pool-sequencing data and genetic diversity

After filtering out reads with a mapping quality below 20, pool-sequencing of the new time point for the three Portuguese populations yielded average median coverages of 83× for autosomes (A) and 63× for the X chromosome (X). Among these, PT-OAZ23 exhibited the highest median coverage (89× for A and 67× for X), while PT-CP23 had the lowest (78× for A and 59× for X) (Suppl. material 2: table S1). Samples from the previous time point, as well as the 22 global populations obtained from Olazcuaga et al. (2020) showed median coverages for autosomes and the X chromosome comparable to those reported in the original studies (Suppl. material 2: table S1).

Population diversity estimates, at the autosomes and X chromosome, of the focal populations confirmed the low diversity previously reported for the PT-VM19 population (Fig. 1). This previously observed pattern holds for both autosomes (Fig. 1c, e; $p < 0.05$, Tukey's HSD test, Suppl. material 2: tables S4, S6) and X chromosome (Fig. 1d, f; $p < 0.05$, Tukey's HSD test, Suppl. material 2: tables S5, S7). In the more recent sample of this population, PT-VM23, diversity has increased to levels comparable to those of the other two Portuguese populations. The only significant differences in Watterson's *theta* were observed between PT-VM23 and PT-OAZ21 in the autosomes and between PT-VM23 and PT-CP21 in the X chromosome ($p < 0.05$, Tukey's HSD test, Suppl. material 2: tables S6, S7). Additionally, in the newer samples, genetic diversity (θ and π) of the PT-CP21 and PT-OAZ21 populations has remained stable ($p > 0.05$, Tukey's HSD test), with a slight increase in diversity at the X chromosome, although this increase is only significant in the PT-CP population (Fig. 1d, f; $p < 0.05$, Tukey's HSD test, Suppl. material 2: tables S5, S7).

Genetic differentiation and population structure

In the focal populations, pairwise F_{ST} estimates between sampling years indicate a general lack of highly differentiated regions (Suppl. material 1: fig. S3). The exception is the comparison between PT-VM19 and PT-VM23, where overall differentiation is higher, with distinct regions showing increased divergence (Suppl. material 1: fig. S3c). Nevertheless, isolated peaks of differentiation are also observed across all chromosomes in the PT-CP and PT-OAZ populations (Suppl. material 1: fig. S3a, b).

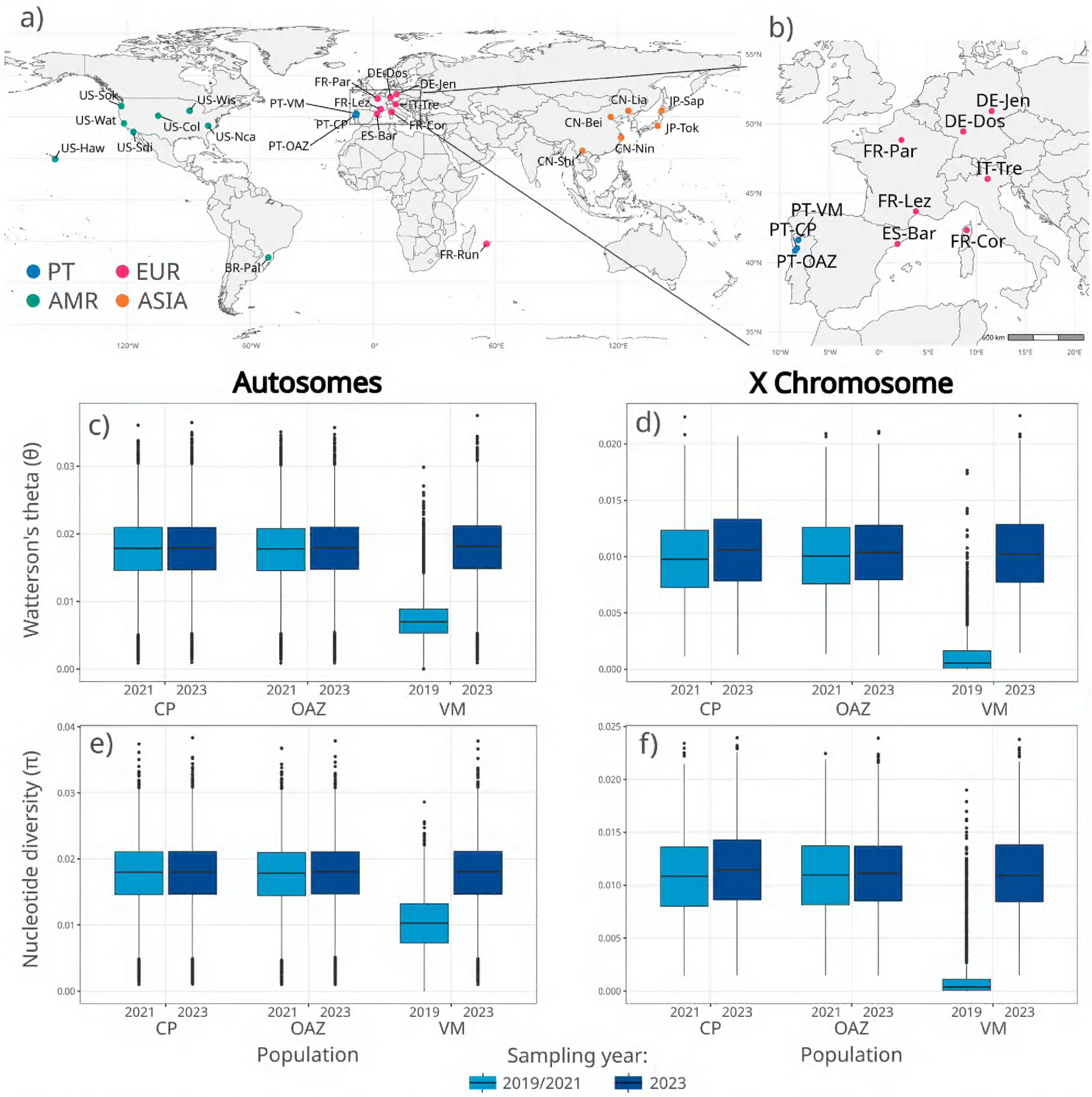

Figure 1. Geographic location of the *Drosophila suzukii* populations analyzed here and genetic diversity of the Portuguese populations. **a.** Global distribution of the populations, including the 22 population from Olazcuaga et al. (2020) and the Portuguese populations from Sario et al. (2024) and this study. Dots are colored according to their origin (teal: America; orange: Asia; pink: Europe and blue: Portuguese (PT) focal populations); **b.** Close-up view of European and Portuguese sampling locations; **c–f.** Measures of genetic diversity across autosomes and the X chromosome for the three focal Portuguese populations (CP, OAZ, and VM) sampled at two time points (2019/2021 and 2023). Watterson's theta (θ) for autosomes (**c**) and the X chromosome (**d**), and nucleotide diversity (π) for autosomes (**e**) and the X chromosome (**f**). Sampling time is indicated by color, with older samples shown in light blue and more recent samples in dark blue.

When considering the world-wide dataset, Pairwise F_{ST} estimates confirmed the previously detected large differentiation between PT-VM19 and the other populations (Fig. 2a). The heatmap also shows that Pairwise F_{ST} between American populations tends to be higher than F_{ST} between European or Asian populations. The scaled covariance matrix of population allele frequencies (Ω) grouped the 28 populations into two distinct clusters - one comprising the Asian and American populations, and the other consisting of the European populations,

including the Portuguese populations (Fig. 2c). Confirming previous results, in the autosomal analysis the Portuguese populations clustered with those from France (FR-Cor), Spain (ES-Bar), and Italy (IT-Tre). Notably, the PT-VM19 population formed a separate branch within the European cluster, while the PT-VM23 population clustered with the remaining Portuguese populations

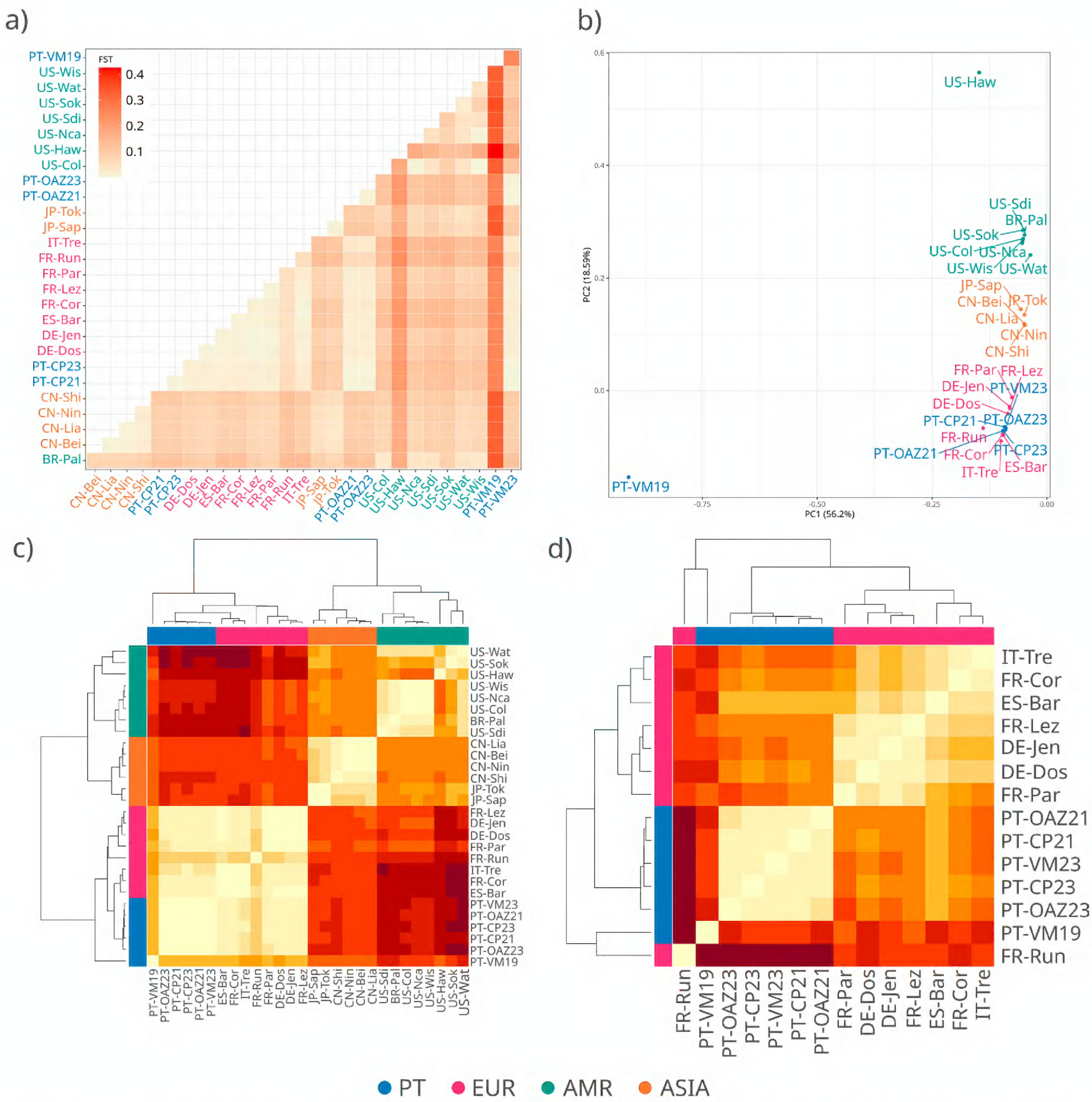

Figure 2. Genetic relationships among *Drosophila suzukii* populations. **a.** Pairwise F_{ST} heatmap with darker shades of red indicating higher genetic differentiation, while lighter shades represent lower differentiation; **b.** Principal Component Analysis (PCA) plot showing the genetic relationships among populations. PC1 explains 56.2% of the variation, while PC2 accounts for 18.59%; **c.** Heatmap of genetic differentiation among *D. suzukii* populations based on the scaled covariance matrix (Ω) estimated using BayPass. The hierarchical clustering dendrogram shows population relationships; **d.** Heatmap of genetic differentiation among European *Drosophila suzukii* populations based on the scaled covariance matrix (Ω) estimated using BayPass. Populations are consistently color-coded by their geographic origin: American populations in teal, Asian populations in orange, European populations in pink, and Portuguese focal populations in blue. In panels (**a**), (**c**), and (**d**), sampling sites are shown along both axes, with Portuguese populations (PT) included alongside populations from the United States (US), Japan (JP), France (FR), Germany (DE), Spain (ES), Italy (IT), China (CN), and Brazil (BR).

(Fig. 2c). The singular-value decomposition of the matrix Ω (SVD of Ω) more clearly highlights the separation of the PT-VM19 population from the remaining populations, as PC1 separates this population from all other populations (Fig. 2b). Only in PC2 do we observe the expected geographical separation, with populations clustering by continent. As seen in Sario et al. (2024), and in the matrix Ω heatmap, the other Portuguese populations, including PT-VM23, were grouped with the rest of the European populations (Fig. 2b).

The scaled covariance matrix of population allele frequencies (Ω) computed when considering only the European populations grouped the populations into two broad clusters, one containing the Portuguese populations and another with the remaining European populations (Fig. 2d). Note that the PT-VM19 population (probably due to its low diversity) and the FR-Run population (probably due to its geographical location) cluster apart from the remaining populations. Interestingly, all other Portuguese populations clustered by sampling year rather than by geographic location, except for PT-OAZ23 (Fig. 2d). At an even finer-scale, when considering only the Portuguese populations, the Ω matrix clearly separates the PT-VM19 population from all other focal populations (Suppl. material 1: fig. S4). This focus on the Portuguese populations confirmed that the populations tend to group by sampling year and that the PT-OAZ23 population forms a separate branch within the Portuguese cluster (Suppl. material 1: fig. S4).

In the TreeMix analysis, results were broadly concordant with the previous analysis. Although our analysis suggests that m = 6 is the optimum number of migration edges, here we present results with m = 3 to facilitate comparisons with previous work. Our results confirm that, for m = 3, no migration edges involve Portuguese populations, with all edges affecting American and Asian populations (e.g. US-Haw and the Chinese populations; Fig. 3a). TreeMix placed the Portuguese focal populations in the European clade, grouping them with the other South European populations (ES-Bar, FR-Cor, and IT-Tre). Confirming previous results, our analysis also detected a long drift branch for the PT-VM19 population, in keeping with its low level of genetic diversity (Fig. 3a).

Allowing for six migration edges resulted in heterogeneous topologies across replicates, with three distinct maximum likelihood topologies inferred. However, in all cases, the inferred admixture edges consistently involved the same populations. Signals involving the American and Asian populations were consistent with those observed when m = 3, while additional edges now connected European populations (Fig. 3b). Specifically, we detected admixture events between Southern European populations and French/German populations, as well as signals involving the Portuguese populations, with an edge connecting PT-CP23 and the node leading to PT-VM19 (Fig. 3b). It is important to note that TreeMix does not explicitly model contemporary migration but rather historical mixing among ancestral lineages; thus, these results are best interpreted as evidence of shared ancestry or past admixture events among these populations. When the TreeMix analysis was restricted to European populations (with JP-Sap included as an outgroup), the relative placement of the Portuguese populations remained unchanged, forming a single clade closely related to other Southern European populations (Suppl. material 1: fig. S5). As before, we detected an admixture signal involving the PT-VM19 lineage, this time connected to the FR-Par population. Interestingly, with m = 3, an additional edge connected the JP-Sap outgroup with the FR-Lez population (Suppl. material 1: fig. S5).

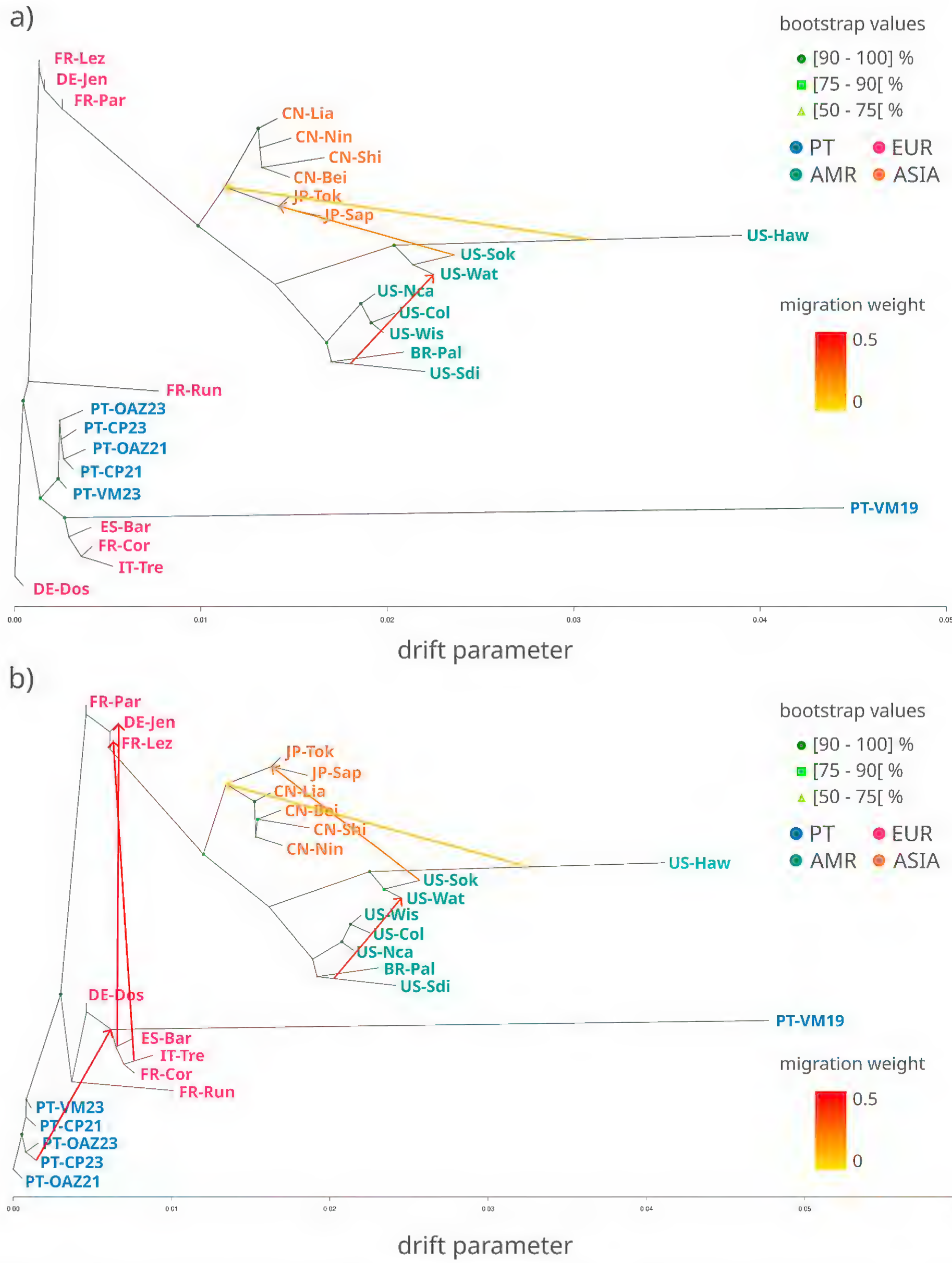

Figure 3. TreeMix analysis showing the inferred population relationships among *D. suzukii* populations, with drift parameter values along the x-axis. **a.** TreeMix tree of relationships with m = 3; **b.** TreeMix tree of relationships with m = 6. Bootstrap values are represented by different symbols: dark green circles for [90–100]% support, light green squares for [75–90]% and light green triangles for [50–75]%. Populations are color-coded by their geographic origin: American populations in teal, Asian populations in orange, European populations in pink, and Portuguese focal populations in blue. The TreeMix analyses presented here were conducted without specifying an outgroup population; therefore, the resulting trees are unrooted.

Genomic regions under selection

Pool-hmm analysis identified several genomic regions classified under the hidden state 'selection' in the Portuguese populations. Each population was analyzed independently, and the numbers reported for each location summarize the windows detected across the different sampling years: 212 in the PT-CP populations (PT-CP21 and PT-CP23), 190 in the PT-OAZ populations (PT-OAZ21 and PT-OAZ23), and 584 in the PT-VM populations (PT-VM19 and PT-VM23). Consistent with previous findings (Sario et al. 2024), the high number of genomic windows under selection in the PT-VM populations is primarily driven by PT-VM19, which alone accounts for 462 windows classified under the hidden state 'selection'. This discrepancy is likely due to the notably low levels of diversity in PT-VM19, suggesting a significant deviation from demographic equilibrium and potentially leading to an overestimation of regions classified as under selection.

Considering each region where we detected signatures of selection, together with 5kb on either side, we identified and annotated 1,109 unique genes across Portuguese populations, of which 500 have known functions. However, most of those 1,109 genes are detected in the older samples (Fig. 4a), with a disproportionate amount being detected in the PT-VM19 population (Suppl. material 1: fig. S6c). In contrast to the PT-VM populations, the PT-CP and PT-OAZ populations exhibited a higher number of genes under selection in the more recent samples (Suppl. material 1: fig. S6a, b). Sario et al. (2024) previously identified candidate genes for the Portuguese invasion by using Pool-hmm to detect genomic regions with signatures of selection. Genomic fragments classified under the 'selection' state, particularly those with the strongest signals, were considered likely contributors to adaptation during the invasion. Comparing the genes detected in our analyses with those candidates confirmed the significance of six genes (*LanB1*, *Dpm2*, *Rbp6*, *bab1*, *Muc68Ca*, and *p130CAS*; Fig. 4b).

The intersection of genes inferred to be under selection across the Portuguese populations identified 40 shared genes, of which 9 have known functions (*Sema1a*, *nrm*, *Gpdh1*, *Acer*, *H2.0*, *Uba4*, *hkb*, *l(2)k09848* and *beta-Man*; Fig. 4a). We included most of these genes in our primer design. Other potentially interesting targets are only detected in the PT-CP and PT-OAZ populations (such as *Muc68Ca*), or in the PT-VM population (e.g. *Gr22a*, *Gr22e* and *Dsk*). The analysis of gene ontology (GO) term enrichment in the set of genes inferred under selection across focal populations (1,109 genes) identified terms associated with protein and heparin binding, as well as immune system processes along with molecular function regulator activity and multicellular organismal processes (Suppl. material 2: table S8).

Visualization of the Pool-hmm results revealed a highly heterogeneous distribution of genomic windows classified under the hidden state 'selection' across all chromosomes. However, these windows are notably concentrated towards the chromosome ends and there is a substantial portion of chromosome 3L identified as being under selection (Suppl. material 1: fig. S7). A comparison of genomic windows identified by Pool-hmm with those classified as F_{ST} outliers and outliers in Fisher's exact test—used to assess the significance of allele frequency differences between Portuguese populations—revealed substantial agreement among these approaches. Most genomic windows classified under the hidden state 'selection' by Pool-hmm were associated with either an F_{ST} outlier window or a SNP with a significantly different allele frequency (after Bonferroni adjustment; Suppl. material 1: fig. S7).

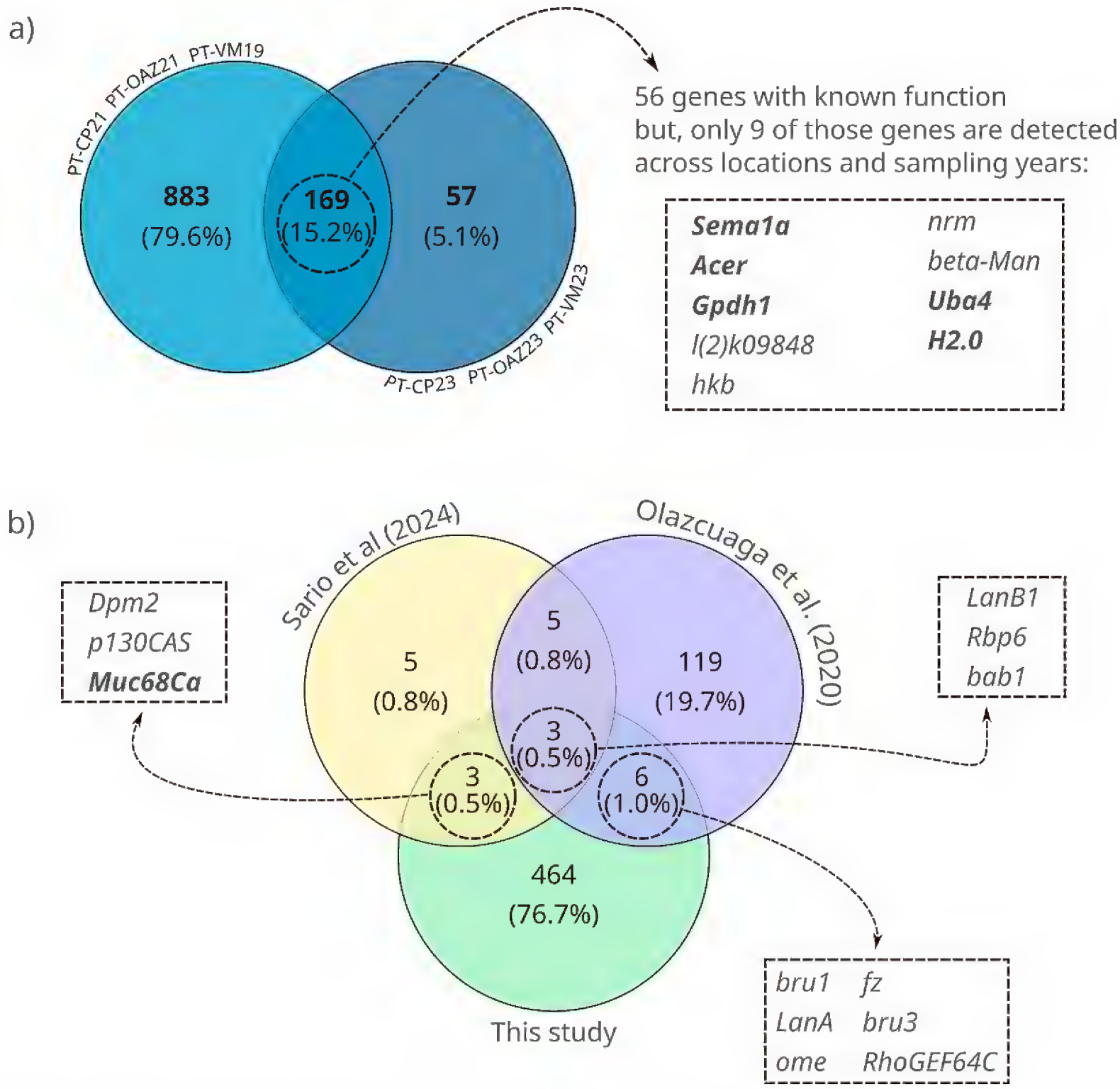

Figure 4. Venn diagrams illustrating the overlap of genes in genomic regions with signatures of selection. **a.** Comparison between the two points in our time-series data, showing the number and proportion of unique and shared genes between earlier (PT-CP21, PT-OAZ21, PT-VM19) and later (PT-CP23, PT-OAZ23, PT-VM23) samples. Each population was analyzed independently, and the diagrams summarize the genes detected at different time points for the same location. We highlight the genes with a known function detected as under selection across locations and sampling years; **b.** Overlap between the list of genes inferred to be under selection in the focal populations in this study, and those reported by Sario et al. (2024) and by Olazcuaga et al. (2020). In both panels, numbers within each section indicate the count and percentage of genes located in genomic regions showing signatures of selection. Genes highlighted in bold were selected for primer design and validation.

Genetic resources to track the Portuguese invasion

All primer pairs successfully amplified the expected regions. However, sequencing quality varied. For *Gpdh1* (G), high background noise in the female sample interfered with base calling, and a secondary sequence was observed in the male sample. The final SNP in this gene and one SNP in the female sample could not be validated. For *Sema1a* (S1), the presence of a microsatellite may have reduced sequencing quality, though it did not interfere with SNP identification. Sequences for the remaining markers (*Acer*, *H2.0*, *Uba4*, and *Muc68Ca*) showed good quality and were successfully validated. Alignment of the corrected sequences with the reference confirmed the presence of the target SNPs, excluding primer regions. The validated SNPs and their flanking sequences provide a robust set of markers for future monitoring efforts, with the potential to detect temporal and spatial changes in adaptive allele frequencies.

Discussion

Diversity estimates (θ and π) for the Portuguese populations confirmed that PT-VM19 represents a unique case across the sampled populations. Its markedly lower genetic diversity was not observed in any other Portuguese population. This distinct genetic profile is unlikely to result from data quality issues, as read mapping and filtering statistics were comparable across all samples. Although we cannot completely exclude minor batch effects caused by unequal individual contribution, this would be expected to contribute to subtler differences. Notably, a population sampled just 4 years later from the same location (PT-VM23) exhibited diversity levels comparable with the other Portuguese populations analyzed here (Fig. 1c–f). This suggests that the low diversity of the older sample was due to the founder's effect and that the sample was likely taken shortly after the establishment of the population. This aligns with the observed Tajima's D, as the high genome-wide D value for PT-VM at the autosomes (1.67 $\pm$ 1.72) suggests that the founder's effect was probably occurring (Nei et al. 1975; Lewald et al. 2021). Furthermore, the homogenization of diversity levels across Portuguese populations is consistent with findings from German and Italian *D. suzukii* populations, where no significant differences in diversity were observed between populations (Tait et al. 2017; Petermann et al. 2021). Interestingly, this rapid recovery of diversity levels likely reflects the short generation time of *Drosophila suzukii* (~12–15 days; Little et al. 2020), its large population sizes ($N \approx 2.5 \times 10^5$–$2.5 \times 10^6$; Adrion et al. 2014), and the potential contribution of migration. Our TreeMix results with m = 6 seem to agree with this hypothesis, as we detected a migration edge affecting the PT-CP23 population and the node leading to PT-VM19 (Fig. 3b). Although this seems counterintuitive, as PT-VM19 was sampled four years earlier, this result likely reflects methodological limitations of TreeMix, which infers migration from allele frequency covariance without accounting for sampling time. Biologically, this signal may instead represent shared ancestry, gene flow from unsampled intermediate populations, or historical admixture affecting both populations.

Our analysis of population structure using BayPass is congruent with previous results. For the world-wide dataset, there is a clear separation between European and American/Asian populations (Fig. 2a). This is expected as the colonization of these two geographical regions represents two independent demographic events. Within the European cluster, we confirmed that all Portuguese populations, except for PT-VM19, clustered together and were close to the remaining South European population (IT-Tre, ES-Bar and FR-Cor). Clustering of the PT-VM23 population in the Portuguese group suggests that PT-VM19 does not represent a different colonization history from the other Portuguese populations (Fig. 2a). When the analysis was focused on the European populations, we detected a clear separation of the Portuguese populations from the remaining ones (Fig. 2c). This suggests some degree of isolation between Portuguese populations and the other European populations, possibly due to the invasion founder events. Interestingly, when the analysis was restricted to Portuguese populations, we observed clustering by sampling year rather than by location, suggesting that these populations may have evolved along a common trajectory (Suppl. material 1: fig. S4). However, an alternative explanation is that some local populations, such as PT-VM19, may have gone extinct and were subsequently replaced by new migrants from other sources. These two processes—shared evolutionary trajectories and recurrent migration—are not mutually exclusive and could both contribute to the observed temporal clustering.

To understand the evolutionary history of the Portuguese populations, we estimated the relationship between *D. suzukii* populations using TreeMix. Previous results, based on the earlier sampling time point, were unable to fully explain the genetic patterns of the PT-VM19 population, suggesting some level of heterogeneity in the Portuguese invasion route (Sario et al. 2024). Here, we show that the same population, sampled 4 years later, is placed with the remaining Portuguese populations (Fig. 3a), suggesting that there is no difference in the invasion route of this population. Interestingly, in this analysis, the Portuguese populations do not cluster strictly by sampling year or locality. This pattern likely reflects key biological characteristics of *D. suzukii*, including large effective population sizes, high dispersal capacity, and recurrent introductions and admixture. These features can erode genetic differentiation across space and time, particularly among populations sampled within close geographic and temporal proximity. Consequently, TreeMix clustering likely captures shared ancestry and ongoing connectivity rather than temporal or spatial structure. Additionally, since the TreeMix graphs presented here are unrooted (no outgroup was specified, except for the analysis shown in Suppl. material 1: fig. S5), branch orientation should not be interpreted as reflecting clustering order, but rather relative levels of genetic drift between populations. Our results with the worldwide dataset from Olazcuaga et al. (2020) confirm the previously suggested invasion route for Portugal (Sario et al. 2024). The close genetic proximity of Portuguese populations, including PT-VM23, to those from Italy, southern France, and Spain suggests a southern European invasion route along the Mediterranean coast. This interpretation is consistent with the earliest record of *D. suzukii* in Portugal, documenting adults in a raspberry greenhouse on the Alentejo coast, in southern Portugal (OEPP/EPPO 2012).

Biological invasions are often associated with a variety of evolutionary adaptive processes, making it challenging to differentiate between adaptations specific to a particular region and those related to the invasion process itself (Pérez et al. 2006; Estoup et al. 2016). Comparing our results with previous studies on the global invasion of *D. suzukii* (Olazcuaga et al. 2020) and its establishment in Portugal (Sario et al. 2024), we identified 12 genetic loci as potential shared targets of selection. However, out of all those genes (*LanB1*, *Dpm2*, *Rbp6*, *bab1*, *p130CAS*, *Muc68Ca*, *bru1*, *LanA*, *ome*, *fz*, *bru3*, *RhoGEF64C*), only *Muc68Ca* is consistently detected across all other populations when PT-VM19 is excluded from the comparison. In *Drosophila melanogaster*, mucin genes such as *Muc68E* have been shown to influence adult size, starvation tolerance, and cold recovery (Reis et al. 2016), while *Muc68Ca* may play a role in heat shock response (Jensen et al. 2008). Given that *Muc68Ca* is expressed in larval/embryonic midgut and salivary gland and predicted to be an extracellular matrix component (Vonesch et al. 2016), it may similarly contribute to gut barrier integrity, nutrient absorption, or protection from environmental stress (Syed et al. 2008). Selection on *Muc68Ca* in Portuguese populations may thus reflect adaptation to abiotic stresses (e.g. thermal extremes) or biotic factors (diet, pathogens) that differ in invaded habitats. Functional studies (e.g., expression under stress or knockdown experiments) are required to test this hypothesis. Previous research on the Portuguese populations suggested that the *nAChRalpha7* gene could play a role in local adaptation during the Portuguese invasion, potentially contributing to insecticide resistance (Sario et al. 2024). However, our reanalysis of this data, combined with the addition of a second time point, did not confirm the importance of this gene (Fig. 4). In fact, polymorphic sites located within the *nAChRalpha7* gene region and classified as

having a predicted hidden state of "Selection" by pool-hmm were detected only in the PT-VM19 population, suggesting that this signal may be a false positive.

This study highlights the value of time-series data in distinguishing between demographic effects and signals of selection. By analyzing the focal Portuguese populations at two distinct time points, we were able to more precisely identify candidate genetic loci that may have been targets of selection during the invasion of *D. suzukii* in Portugal. This was achieved by focusing our analysis on genomic windows showing the strongest signals of selection, consistently detected across all populations and both time points. This analysis showed that nine genes with known functions are consistently predicted to be under selection across the focal populations. However, as with any scan based on allele frequency differentiation, we cannot exclude the possibility that some of these signals arise from linkage disequilibrium with nearby causal loci rather than direct selection acting on the annotated genes themselves. These genes are associated with diverse biological functions, including neural development (*Sema1a* and *nrm*), embryogenesis (*hkb*), metabolism (*Gpdh1*, *Acer*, and *beta-Man*), and chromatin organization (*H2.0*). Among the genes consistently identified as being under selection in all focal Portuguese populations, *Uba4* stands out by having functions that closely reflect the terms identified by our GO term enrichment analysis. This gene plays a critical role in protein quality control by tagging proteins for degradation, a process that influences several important signaling pathways (Khoshnood et al. 2015). While definitive evidence will require functional validation, the consistent temporal signals we observe point to these genes as plausible contributors to local adaptation. Given the significant economic impact of *D. suzukii* (Bolda et al. 2010; Yeh et al. 2020), continued monitoring of its invasion dynamics in Portugal is essential. Future efforts should focus on tracking the temporal patterns of putatively adaptive alleles to better understand the genomic changes driving invasion success. Confirming consistent shifts in allele frequencies over extended time periods will help identify key variants under selection. To support this, we have developed and validated a set of SNP markers targeting genes with the strongest signatures of selection (*Gpdh1*, *Sema1a*, *Acer*, *H2.0*, *Uba4*, and *Muc68Ca*). This marker set provides a valuable tool for future surveillance, enabling the detection of temporal and spatial shifts in adaptive genetic variation within Portuguese populations of *D. suzukii*.

Conclusion

Our time-series analysis of the *D. suzukii* invasion in Portugal, the westernmost edge of the European invasion, supported a Mediterranean invasion route through Europe, and clarified that the low genetic diversity previously estimated in PT-VM19 was a consequence of a founder effect rather than a distinct invasion history. Our data showed a rapid recovery of genetic diversity in this population. While we identified signatures of positive selection along the genome, much of the overlap with genes previously linked to the species' invasive capacity is attributable to PT-VM19. Since many of the selection signals exclusive to this population were likely false positives, the absence of such overlap in other Portuguese populations suggests that the invasion was characterized by unique selection signatures, potentially reflecting local adaptation. To facilitate future monitoring, we developed a set of genetic markers to track adaptive variation in Portuguese populations. Overall,

our findings underscore the complexity of rapid species invasions, the role of natural selection in shaping invasive success, and how temporal genomic analyses are key to unravel these demographic and adaptive dynamics.

Acknowledgements

The authors would like to express their gratitude to the producers who generously agreed to take part in this study and granted access to their orchards.

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Use of AI

No use of AI was reported.

Funding

This work was supported by project DrosuGreen (grant number PTDC/ASP-PLA/4477/2020), funded by Portuguese Fundação para a Ciência e Tecnologia (FCT). Additional support was obtained from the European Union's Horizon 2020 Research and Innovation Programme under the Grant Agreement Number 857251. The work was co-funded by the European Union through the NORTE 2030 Regional Programme under the NORTE2030-FEDER-02130700 project. We thank the National Network for Advanced Computing (RNCA) and INCD (https://incd.pt/) for access and use of their computing infrastructure, funded by FCT (2025.00006.CPCA.A2).

Author contributions

Conceptualization: JPM, CS, JM-F; Supervision: JPM, CS, JM-F; Funding Acquisition: CS, JM-F; Project administration: CS, JM-F; Investigation: JC, JPM, SS, LF; Methodology: JC, JPM, LF, JM-F; Data curation: JC, JPM; Formal analysis: JC, JPM, SS, JM-F; Validation: JC, JPM; Visualization: JC, JPM; Writing – original draft: JC, JPM, JM-F; Writing – review and edit: JC, JPM, SS, LF, CS, JM-F.

Author ORCIDs

João Carvalho https://orcid.org/0000-0002-1728-0075
João Pedro Marques https://orcid.org/0000-0001-9834-1361
Sara Sario https://orcid.org/0000-0002-3695-9227
Liliana Farelo https://orcid.org/0000-0003-4985-3592
Conceição Santos https://orcid.org/0000-0003-4129-6381
José Melo-Ferreira https://orcid.org/0000-0003-4473-1908

Data availability

Whole genome sequencing data for the original Portuguese samples (PT-VM19, PT-CP21 and PT-OAZ21) is available at NCBI under the BioProject PRJNA1081763 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1081763). Data for the newer samples in our time-series (PT-VM23, PT-CP23

and PT-OAZ23) is available at NCBI under the BioProject PRJNA1289399 (http://www.ncbi.nlm.nih.gov/bioproject/1289399). The sequences and alignments used for SNP validation are available on Zenodo at https://doi.org/10.5281/zenodo.17370752.

References

Adrion JR, Kousathanas A, Pascual M, Burrack HJ, Haddad NM, Bergland AO, Machado H, Sackton TB, Schlenke TA, Watada M, Wegmann D, Singh ND (2014) *Drosophila suzukii*: The genetic footprint of a recent, worldwide invasion. Molecular Biology and Evolution 31: 3148–3163. https://doi.org/10.1093/molbev/msu246

Andolfatto P (2001) Adaptive hitchhiking effects on genome variability. Current Opinion in Genetics & Development 11: 635–641. https://doi.org/10.1016/s0959-437x(00)00246-x

Barnett DW, Garrison EK, Quinlan AR, Strömberg MP, Marth GT (2011) BamTools: A C++ API and toolkit for analyzing and managing BAM files. Bioinformatics (Oxford, England) 27: 1691–1692. https://doi.org/10.1093/bioinformatics/btr174

Barrett SCH (2015) Foundations of invasion genetics: The Baker and Stebbins legacy. Molecular Ecology 24: 1927–1941. https://doi.org/10.1111/mec.13014

Barton NH (2000) Genetic hitchhiking. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 355: 1553–1562. https://doi.org/10.1098/rstb.2000.0716

Bergland AO, Behrman EL, O'Brien KR, Schmidt PS, Petrov DA (2014) Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in *Drosophila*. PLOS Genetics 10: e1004775. https://doi.org/10.1371/journal.pgen.1004775

Boitard S, Kofler R, Françoise P, Robelin D, Schlötterer C, Futschik A (2013) Pool-hmm: A Python program for estimating the allele frequency spectrum and detecting selective sweeps from next generation sequencing of pooled samples. Molecular Ecology Resources 13: 337–340. https://doi.org/10.1111/1755-0998.12063

Bolda MP, Goodhue RE, Zalom FG (2010) Spotted wing drosophila: Potential economic impact of a newly established pest. Agricultural and Resource Economics Update 13: 5–8.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics (Oxford, England) 30: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Boughdad A, Haddi K, El Bouazzati A, Nassiri A, Tahiri A, El Anbri C, Eddaya T, Zaid A, Biondi A (2020) First record of the invasive spotted wing drosophila infesting berry crops in Africa. Journal of Pest Science 94: 261–271. https://doi.org/10.1007/s10340-020-01280-0

Calabria G, Máca J, Bächli G, Serra L, Pascual M (2010) First records of the potential pest species *Drosophila suzukii* (Diptera: Drosophilidae) in Europe. Journal of Applied Entomology 136: 139–147. https://doi.org/10.1111/j.1439-0418.2010.01583.x

Cini A, Anfora G, Escudero-Colomar LA, Grassi A, Santosuosso U, Seljak G, Papini A (2014) Tracking the invasion of the alien fruit pest *Drosophila suzukii* in Europe. Journal of Pest Science 87(4): 559–566. https://doi.org/10.1007/s10340-014-0617-z

Colautti RI, Barrett SCH (2013) Rapid adaptation to climate facilitates range expansion of an invasive plant. Science 342: 364–366. https://doi.org/10.1126/science.1242121

Colautti RI, Lau JA (2016) Contemporary evolution during invasion: evidence for differentiation, natural selection, and local adaptation. Invasion Genetics: The Baker and Stebbins Legacy. Wiley, 101–121. https://doi.org/10.1002/9781119072799.ch6

Daetwyler HD, Capitan A, Pausch H, Stothard P, van Binsbergen R, Brøndum RF, Liao X, Djari A, Rodriguez SC, Grohs C, Esquerré D, Bouchez O, Rossignol M-N, Klopp C, Rocha D, Fritz S, Eggen A, Bowman PJ, Coote D, Chamberlain AJ, Anderson C, VanTassell CP, Hulsegge I, Goddard ME, Guldbrandtsen B, Lund MS, Veerkamp RF, Boichard DA, Fries R, Hayes BJ (2014)

Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. Nature Genetics 46: 858–865. https://doi.org/10.1038/ng.3034

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, 1000 Genomes Project Analysis Group (2011) The variant call format and VCFtools. Bioinformatics (Oxford, England) 27: 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

De Ros G, Grassi A, Pantezzi T (2020) Recent Trends in the Economic Impact of *Drosophila suzukii*. *Drosophila suzukii* Management. Springer International Publishing, Cham, 11–27. https://doi.org/10.1007/978-3-030-62692-1_2

Dlugosch KM, Anderson SR, Braasch J, Cang FA, Gillette HD (2016) The devil is in the details: genetic variation in introduced populations and its contributions to invasion. Invasion Genetics: The Baker and Stebbins Legacy. Wiley, 232–251. https://doi.org/10.1002/9781119072799.ch14

Estoup A, Ravigné V, Hufbauer R, Vitalis R, Gautier M, Facon B (2016) Is there a genetic paradox of biological invasion? Annual Review of Ecology, Evolution, and Systematics 47: 51–72. https://doi.org/10.1146/annurev-ecolsys-121415-032116

Facon B, Hufbauer RA, Tayeh A, Loiseau A, Lombaert E, Vitalis R, Guillemaud T, Lundgren JG, Estoup A (2011) Inbreeding depression is purged in the invasive insect *Harmonia axyridis*. Current Biology: CB 21: 424–427. https://doi.org/10.1016/j.cub.2011.01.068

Fisher RA, Ford EB (1947) The spread of a gene in natural conditions in a colony of the moth *Panaxia dominula* L. Heredity 1: 143–174. https://doi.org/10.1038/hdy.1947.11

Fitak RR (2021) OptM: Estimating the optimal number of migration edges on population trees using Treemix. Biology Methods and Protocols 6. https://doi.org/10.1093/biomethods/bpab017

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A bayesian perspective. Genetics 180: 977–993. https://doi.org/10.1534/genetics.108.092221

Fraimout A, Debat V, Fellous S, Hufbauer RA, Foucaud J, Pudlo P, Marin J-M, Price DK, Cattel J, Chen X, Deprá M, Duyck PF, Guedot C, Kenis M, Kimura MT, Loeb G, Loiseau A, Martinez-Sañudo I, Pascual M, Richmond MP, Shearer P, Singh N, Tamura K, Xuéreb A, Zhang J, Estoup A (2017) Deciphering the routes of invasion of *Drosophila suzukii* by means of ABC random forest. Molecular Biology and Evolution 34: 980–996. https://doi.org/10.1093/molbev/msx050

Futschik A, Schlötterer C (2010) The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. Genetics 186: 207–218. https://doi.org/10.1534/genetics.110.114397

Garcia FRM, Lasa R, Funes CF, Buzzetti K (2022) *Drosophila suzukii* management in Latin America: Current status and perspectives. Journal of Economic Entomology 115: 1008–1023. https://doi.org/10.1093/jee/toac052

Gaur PM, Samineni S, Thudi M, Tripathi S, Sajja SB, Jayalakshmi V, Mannur DM, Vijayakumar AG, Rao NVPRG, Ojiewo C, Fikre A, Kimurto P, Kileo RO, Girma N, Chaturvedi SK, Varshney RK, Dixit GP (2018) Integrated breeding approaches for improving drought and heat adaptation in chickpea (*Cicer arietinum* L.). Plant Breeding 138: 389–400. https://doi.org/10.1111/pbr.12641

Gautier M (2015) Genome-Wide scan for adaptive divergence and association with population-specific covariates. Genetics 201: 1555–1579. https://doi.org/10.1534/genetics.115.181453

Gautier M, Vitalis R, Flori L, Estoup A (2021) f-Statistics estimation and admixture graph construction with Pool-Seq or allele count data using the R package poolfstat. Molecular Ecology Resources 22: 1394–1416. https://doi.org/10.1111/1755-0998.13557

Hauser M (2011) A historic account of the invasion of *Drosophila suzukii* (Matsumura) (Diptera: Drosophilidae) in the continental United States, with remarks on their identification. Pest Management Science 67: 1352–1357. https://doi.org/10.1002/ps.2265

Hill DL, Carr EJ, Rutishauser T, Moncunill G, Campo JJ, Innocentin S, Mpina M, Nhabomba A, Tumbo A, Jairoce C, Moll HA, van Zelm MC, Dobaño C, Daubenberger C, Linterman MA (2020) Immune system development varies according to age, location, and anemia in African children. Science Translational Medicine 12. https://doi.org/10.1126/scitranslmed.aaw9522

Jenkins VK, Larkin A, Thurmond J (2022) Using flybase: A database of drosophila genes and genetics. Methods in Molecular Biology. Springer US, New York, NY, 1–34. https://doi.org/10.1007/978-1-0716-2541-5_1

Jensen LT, Nielsen MM, Loeschcke V (2008) New candidate genes for heat resistance in *Drosophila melanogaster* are regulated by HSF. Cell Stress & Chaperones 13: 177–182. https://doi.org/10.1007/s12192-008-0020-x

Khoshnood B, Dacklin I, Grabbe C (2015) Urm1: An essential regulator of JNK signaling and oxidative stress in *Drosophila melanogaster*. Cellular and Molecular Life Sciences: CMLS 73: 1939–1954. https://doi.org/10.1007/s00018-015-2121-x

Knapp L, Mazzi D, Finger R (2020) The economic impact of *Drosophila suzukii*: Perceived costs and revenue losses of Swiss cherry, plum and grape growers. Pest Management Science 77: 978–1000. https://doi.org/10.1002/ps.6110

Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK (2012) VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Research 22: 568–576. https://doi.org/10.1101/gr.129684.111

Kofler R, Orozco-terWengel P, De Maio N, Pandey RV, Nolte V, Futschik A, Kosiol C, Schlötterer C (2011) PoPoolation: A toolbox for population genetic analysis of next generation sequencing data from pooled individuals. PLOS ONE 6: e15925. https://doi.org/10.1371/journal.pone.0015925

Lee CE, Gelembiuk GW (2008) Evolutionary origins of invasive populations. Evolutionary Applications 1: 427–448. https://doi.org/10.1111/j.1752-4571.2008.00039.x

Lewald KM, Abrieux A, Wilson DA, Lee Y, Conner WR, Andreazza F, Beers EH, Burrack HJ, Daane KM, Diepenbrock L, Drummond FA, Fanning PD, Gaffney MT, Hesler SP, Ioriatti C, Isaacs R, Little BA, Loeb GM, Miller B, Nava DE, Rendon D, Sial AA, da Silva CSB, Stockton DG, Timmeren SV, Wallingford A, Walton VM, Wang X, Zhao B, Zalom FG, Chiu JC (2021) Population genomics of *Drosophila suzukii* reveal longitudinal population structure and signals of migrations in and out of the continental United States. G3 Genes|Genomes|Genetics 11(12): jkab343. https://doi.org/10.1093/g3journal/jkab343

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics (Oxford, England) 25: 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics (Oxford, England) 25: 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Little CM, Chapman TW, Hillier NK (2020) Plasticity is key to success of *Drosophila suzukii* (Diptera: Drosophilidae) invasion. Journal of Insect Science 20(3): 5. https://doi.org/10.1093/jisesa/ieaa034

Macedo FL, Ragonezi C, Reis F, de Freitas JGR, Lopes DH, Aguiar AMF, Cravo D, de Carvalho MAAP (2023) Prediction of the potential distribution of *Drosophila suzukii* on Madeira Island using the maximum entropy modeling. Agriculture 13: 1764. https://doi.org/10.3390/agriculture13091764

Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, Harney E, Stewardson K, Fernandes D, Novak M, Sirak K, Gamba C, Jones ER, Llamas B, Dryomov S, Pickrell J, Arsuaga JL, de Castro JMB, Carbonell E, Gerritsen F, Khokhlov A, Kuznetsov P, Lozano M, Meller H, Mochalov O, Moiseyev V, Guerra MAR, Roodenberg J, Vergès JM, Krause J, Cooper A, Alt KW, Brown D, Anthony D, Lalueza-Fox C, Haak W, Pinhasi R, Reich D (2015) Genome-wide patterns of selection in 230 ancient Eurasians. Nature 528: 499–503. https://doi.org/10.1038/nature16152

Min J, Gupta M, Desai MM, Weissman DB (2022) Spatial structure alters the site frequency spectrum produced by hitchhiking. Genetics 222. https://doi.org/10.1093/genetics/iyac139

Natterson-Horowitz B, Aktipis A, Fox M, Gluckman PD, Low FM, Mace R, Read A, Turner PE, Blumstein DT (2023) The future of evolutionary medicine: Sparking innovation in biomedicine and public health. Frontiers in Science 1. https://doi.org/10.3389/fsci.2023.997136

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. Evolution; International Journal of Organic Evolution 29: 1. https://doi.org/10.2307/2407137

Nielsen R, Williamson S, Kim Y, Hubisz MJ, Clark AG, Bustamante C (2005) Genomic scans for selective sweeps using SNP data. Genome Research 15: 1566–1575. https://doi.org/10.1101/gr.4252305

Ochocki BM, Miller TEX (2017) Rapid evolution of dispersal ability makes biological invasions faster and more variable. Nature Communications 8: 14315. https://doi.org/10.1038/ncomms14315

OEPP/EPPO (2012) First report of *Drosophila suzukii* in Portugal. EPPO Reporting Service No. 10. Paris, France. https://gd.eppo.int/reporting/article-2415

OEPP/EPPO (2013) PM 7/115 (1) *Drosophila suzukii*. EPPO Bulletin 43: 417–424. https://doi.org/10.1111/epp.12059

Olazcuaga L, Loiseau A, Parrinello H, Paris M, Fraimout A, Guedot C, Diepenbrock LM, Kenis M, Zhang J, Chen X, Borowiec N, Facon B, Vogt H, Price DK, Vogel H, Prud'homme B, Estoup A, Gautier M (2020) A whole-genome scan for association with invasion success in the fruit fly *Drosophila suzukii* using contrasts of allele frequencies corrected for population structure. Molecular Biology and Evolution 37: 2369–2385. https://doi.org/10.1093/molbev/msaa098

Pavlidis P, Živković D, Stamatakis A, Alachiotis N (2013) SweeD: Likelihood-Based detection of selective sweeps in thousands of genomes. Molecular Biology and Evolution 30: 2224–2234. https://doi.org/10.1093/molbev/mst112

Pérez JE, Nirchio M, Alfonsi C, Muñoz C (2006) The biology of invasions: The genetic adaptation paradox. Biological Invasions 8: 1115–1121. https://doi.org/10.1007/s10530-005-8281-0

Petermann S, Otto S, Eichner G, Schetelig MF (2021) Spatial and temporal genetic variation of *Drosophila suzukii* in Germany. Journal of Pest Science 94: 1291–1305. https://doi.org/10.1007/s10340-021-01356-5

Pickrell J, Pritchard J (2012) Inference of population splits and mixtures from genome-wide allele frequency data. Nature Precedings. https://doi.org/10.1038/npre.2012.6956.1

Prentis PJ, Wilson JRU, Dormontt EE, Richardson DM, Lowe AJ (2008) Adaptive evolution in invasive species. Trends in Plant Science 13: 288–294. https://doi.org/10.1016/j.tplants.2008.03.004

Ratiu AC, Ionascu A, Constantin ND (2024) The impact of Oxford nanopore technologies based methodologies on the genome sequencing and assembly of Romanian strains of *Drosophila suzukii*. Insects 16(1): 2. https://doi.org/10.3390/insects16010002

Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J (2019) g:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Research 47: W191–W198. https://doi.org/10.1093/nar/gkz369

Rego C, Aguiar AF, Cravo D, Boieiro M (2017) Invasive fruit flies (Diptera: Drosophilidae) meet in a biodiversity hotspot. Journal of Entomological Research Society 19: 61–69. https://doi.org/10.5281/zenodo.10764552

Reis M, Silva AC, Vieira CP, Vieira J (2016) The *Drosophila melanogaster* Muc68E Mucin Gene Influences Adult Size, Starvation Tolerance, and Cold Recovery. G3 Genes|Genomes|Genetics 6: 1841–1851. https://doi.org/10.1534/g3.116.029934

Reyes JA, Lira-Noriega A (2020) Current and future global potential distribution of the fruit fly *Drosophila suzukii* (Diptera: Drosophilidae). Canadian Entomologist 152: 587–599. https://doi.org/10.4039/tce.2020.3

Reznick DN, Losos J, Travis J (2018) From low to high gear: There has been a paradigm shift in our understanding of evolution. Ecology Letters 22: 233–244. https://doi.org/10.1111/ele.13189

Rota-Stabelli O, Blaxter M, Anfora G (2013) *Drosophila suzukii*. Current Biology: CB 23(1): R8–R9. https://doi.org/10.1016/j.cub.2012.11.021

Sario S, Marques JP, Farelo L, Afonso S, Santos C, Melo-Ferreira J (2024) Dissecting the invasion history of Spotted-Wing drosophila (*Drosophila suzukii*) in Portugal using genomic data. BMC Genomics 25. https://doi.org/10.1186/s12864-024-10739-8

Schlötterer C, Tobler R, Kofler R, Nolte V (2014) Sequencing pools of individuals — Mining genome-wide polymorphism data without big funding. Nature Reviews Genetics 15: 749–763. https://doi.org/10.1038/nrg3803

Schneider M, Ballvora A, Léon J (2024) Deep genotyping reveals specific adaptation footprints of conventional and organic farming in barley populations—An evolutionary plant breeding approach. Agronomy for Sustainable Development 44. https://doi.org/10.1007/s13593-024-00962-8

Shearer PW, West JD, Walton VM, Brown PH, Svetec N, Chiu JC (2016) Seasonal cues induce phenotypic plasticity of *Drosophila suzukii* to enhance winter survival. BMC Ecology 16. https://doi.org/10.1186/s12898-016-0070-3

Smith JM, Haigh J (1974) The hitch-hiking effect of a favourable gene. Genetics Research 23: 23–35. https://doi.org/10.1017/s0016672300014634

Stephan W, Song YS, Langley CH (2006) The hitchhiking effect on linkage disequilibrium between linked neutral loci. Genetics 172: 2647–2663. https://doi.org/10.1534/genetics.105.050179

Syed ZA, Härd T, Uv A, van Dijk-Härd IF (2008) A potential role for *Drosophila* mucins in development and physiology. PLOS ONE 3(8): e3041. https://doi.org/10.1371/journal.pone.0003041

Tait G, Vezzulli S, Sassù F, Antonini G, Biondi A, Baser N, Sollai G, Cini A, Tonina L, Ometto L, Anfora G (2017) Genetic variability in Italian populations of *Drosophila suzukii*. BMC Genetics 18. https://doi.org/10.1186/s12863-017-0558-7

Taus T, Futschik A, Schlötterer C (2017) Quantifying selection with pool-seq time series data. Molecular Biology and Evolution 34: 3023–3034. https://doi.org/10.1093/molbev/msx225

Vonesch SC, Lamparter D, Mackay TF, Bergmann S, Hafen E (2016) Genome-wide analysis reveals novel regulators of growth in *Drosophila melanogaster*. PLOS Genetics 12(1): e1005616. https://doi.org/10.1371/journal.pgen.1005616

Walsh DB, Bolda MP, Goodhue RE, Dreves AJ, Lee J, Bruck DJ, Walton VM, O'Neal SD, Zalom FG (2011) *Drosophila suzukii* (Diptera: Drosophilidae): Invasive pest of ripening soft fruit expanding its geographic range and damage potential. Journal of Integrated Pest Management 2: G1–G7. https://doi.org/10.1603/ipm10010

Walton VM, Burrack HJ, Dalton DT, Isaacs R, Wiman N, Ioriatti C (2016) Past, present and future of *Drosophila suzukii*: Distribution, impact and management in United States berry fruits. Acta Horticulturae 87–94. https://doi.org/10.17660/actahortic.2016.1117.16

Wright S (1948) On the roles of directed and random changes in gene frequency in the genetics of populations. Evolution; International Journal of Organic Evolution 2: 279. https://doi.org/10.2307/2405519

Yeh DA, Drummond FA, Gómez MI, Fan X (2020) The economic impacts and management of spotted wing drosophila (*Drosophila suzukii*): The case of wild blueberries in maine. Journal of Economic Entomology 113: 1262–1269. https://doi.org/10.1093/jee/toz360

Supplementary material 1

Supplementary figures

Authors: João Carvalho, João Pedro Marques, Sara Sario, Liliana Farelo, Conceição Santos, José Melo-Ferreira

Data type: docx

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/neobiota.107.176004.suppl1

Supplementary material 2

Supplementary tables

Authors: João Carvalho, João Pedro Marques, Sara Sario, Liliana Farelo, Conceição Santos, José Melo-Ferreira

Data type: xlsx

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/neobiota.107.176004.suppl2